Geogenic signatures detectable in topsoils of urban and rural domains in the London region, UK, using parent material classified data

J. D. Appleton, C. C. Johnson, E.L. Ander, and D.M.A. Flight

British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

* Corresponding author: J. D. Appleton, British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

Tel: +44 (0) 115 936 3207

Fax: +44 (0) 115 936 3200

E-mail address: jda@bgs.ac.uk (J. D. Appleton)

Abstract

Systematic mapping of the chemical environment of urban areas from around the world have shown varying degrees of control of element distributions by the underlying parent material (PM). The purpose of the study reported here is to assess whether geogenic signatures that dominate soil chemistry in rural domains of Eastern England and which are not strongly impacted by human activities, can also be detected in the London urban region. A PM soil chemistry mapping method is used to determine the spatial variation of topsoil chemistry data in London and the surrounding rural areas. Analysis of variance (ANOVA) of the soil data for the London region indicates that 26-33% of the variance of Al, Ce, Cs, Ga, K, La, Mg, Mn, Nb, Nd, Rb, Ti, V and Y is explained by soil PM (surface geology), and a slightly lesser proportion (19-25%) of the variance for Ca, Co, Fe, I, Ni, Sc, Sr and Th. In comparison, soil PM explains only 5% of the variance of Cd. The variance of some other elements appears to be influenced by a mixture of geogenic and anthropogenic controls, including As, Ba, Cr, Cu, Mo, P, Pb, Sb, Se, Sn and Zn for which PM controls 12-16% of the variance. Geogenic soil chemistry patterns observed for the elements strongly influenced by PM in the rural areas surrounding London can be quite clearly followed into and through the London urban area. Spatial patterns of a range of elements primarily controlled by PM have not been destroyed even in a major urban centre with a recorded history dating back over 2,000 years and which has been subjected to extensive urban development, destruction and redevelopment especially during the last 200-300 years.

Keywords: London, rural, urban, parent material, soil chemistry

1. Introduction

Parent material is the primary geogenic control on soil geochemistry in recently glaciated rural landscapes such as the British Isles (Rawlins et al., 2003) and for some chemical elements in some urban areas (Appleton and Adlam, 2012). The strong influence of recent glacial deposits on many element distributions in soil is also seen in the continental scale low density regional geochemical mapping of Europe (Salminen et al. 2005; De Vos et al., 2006; Reimann et al., 2013). Appleton and Adlam (2012) demonstrated that a relatively high proportion of the variance of As, Cr, Fe and K is accounted for by PM, especially in those urban areas where there is a particularly strong chemical contrast between geological units. The variance of Pb accounted for by PM is relatively low (2-9%) in many urban areas, including London (Appleton and Adlam, 2012) since anthropogenic contamination is the major control. If the variance accounted for by PM is low (e.g. <10%) then PM mapping is unlikely to be very informative, whereas when the percentage is 20 or higher, PM exerts a significant control on soil chemistry in urban areas (Appleton and Adlam, 2012). Where this is the case, (i) the parent material (PM) soil chemistry mapping method can be used to portray spatial variation in urban soil chemistry data more accurately than the inverse distance weighted interpolated map data conventionally used in geochemical mapping; (ii) PM mapping allows a better visual impression to be obtained of the geological control on soil chemistry than graduated symbol maps. However, the impact on the PM geochemical maps of uncertainty in the position of mapped geological boundaries, especially where there is a lithological transition between adjacent geological units, is acknowledged. Discrepancies between the PM geochemical maps and the distribution of mapped geological units may in some cases help to improve the geological maps. Systematic mapping of the chemical environment of urban areas from around the world have shown varying degrees of control of element distributions by the underlying PM (Johnson et al., 2011). The relationship been geology and geochemistry is generally explored by comparing spatial distributions, usually

based on interpolated geochemical maps, or looking at element associations defined by factor analysis and attributing likely geogenic relationships. The purpose of the study reported here is to quantify whether PM exerts a significant control on the distribution of a range of lithogenic and potentially harmful elements (PHEs) and to assess whether parent material soil chemistry mapping (Appleton et al., 2008, Appleton and Adlam, 2012) can be used to identify geogenic signatures in the London urban and surrounding rural domains. The term domain is used as described in Johnson et al. (2012). Preliminary assessments of the topsoil chemistry of the predominantly urban Greater London area are available in BGS (2011), Knights and Scheib (2011), Scheib et al. (2011).

2. Materials and methods

2.1 Soil Parent Material

A 1: 50 000 scale surface parent material map (Figure 1) based on a simplified geological classification similar to that developed for radon mapping (Miles and Appleton, 2005) was used for the study. The area is underlain by Cretaceous and Palaeogene bedrocks (Table 1) which in some areas are covered with extensive Quaternary superficial deposits (Table 2). Additional detail, derived from Ellison et al. (2004), BGS (2011) and Royse et al. (2012), is available in the Supplementary material.

2.2 Soil chemistry data

The BGS Geochemical Baseline Survey of the Environment (G-BASE) regional topsoil samples were collected at a depth of c. 5–20 cm and a sampling density of 1 sample per 2 km^2 (Johnson et al., 2005). In the London urban area, soil samples were collected from open ground on a 500 m grid at a density of approximately 4 samples per km^2 (BGS, 2011; Flight and Scheib, 2011). At each site, composite samples, based on 5 sub-samples taken at the

centre and four corners of a 20 m square were collected from the topsoil (5–20 cm depth). Forty eight chemical elements were determined in the <2 mm size fraction of the topsoils using X-ray fluorescence spectrometry (XRFS), together with loss on ignition (LOI at 450°C) and pH in all the urban samples and 50% of the rural samples. Sample preparation, analytical methods, and quality control procedures are described in Allen et al. (2011) and Johnson (2011).

2.3 Analysis of variance and cluster analysis

One way analysis of variance (ANOVA) in MINITAB® 15 was used to calculate the proportion of the variation of element concentrations explained by soil parent material (surface geology) using the sum of squares between group means. The F-statistic was utilised to measure significance level (p-value) and the fraction of the variation related to parent material was always significant statistically ($p < 0.0001$). ANOVA statistics assume that population distributions are normal for each group, variances are equal for all groups, and also that observations are randomly and independently representative of the populations (Reimann et al., 2008). All these conditions are unlikely to be met perfectly in soil geochemical data representing a wide variation of parent materials and the influence of a wide range of environmental and anthropogenic factors. Fortunately, ANOVA is little affected by relatively small or moderate departures from homogeneity of variance. The application of a log-transform in general produces more normal distributions with lower skewness coefficients, especially for elements such as Pb where anthropogenic contamination is a significant factor. Data were log-transformed (Log_e) when the skewness coefficient exceeded 1.0. ANOVA and non-parametric Kruskal-Wallis statistics (not reported here) for soil data grouped using the simplified PM classification have comparable p values.

Cluster analysis of variables in MINITAB® 15 was used to identify groups of elements which share common characteristics and distributions. Ag, Bi, Na, Tl, and W data were removed because of concerns regarding data quality. The number of clusters was determined by examining trial dendrograms and identifying an appropriate similarity level (20) for cutting the exploratory dendrograms. The Ward Linkage method and correlation coefficient (distance measure) were used and the resulting clusters do not vary significantly whether the final partition is based on the similarity level or number of clusters (6 appears to be the most appropriate) nor whether average linkage was used instead of the Ward linkage. Cluster analysis was performed on centred logratio transformed variables (Reimann et al., 2008) in order to ensure approximate symmetry of variable distributions and to remove the effects of closure that characterise major element data.

2.4 Soil chemistry mapping using PM classified data

Soil geochemical data have in some cases large positive skewness coefficients so are transformed by taking natural logarithms. To overcome the bias associated with log-normal data, the geometric mean (GM) and Inverse Distance Weighted GM were used for mapping the spatial variation in element concentrations. Using PM polygons as soil chemistry mapping units, it is possible to estimate element concentrations based on local averages, without significant errors at PM boundaries (Appleton et al., 2008; Appleton and Adlam, 2012). This methodology is generally appropriate in situations and for elements where PM explains a relatively high proportion of the variance, but less so for elements where the proportion of variance explained by PM is low, for example where point source anthropogenic contamination is a major factor, such as Pb.

Parent material classes in this study are based on surface geology. Made ground is not used as a soil PM class, because spatial information on the distribution of made ground in urban areas

in the UK is incomplete. The simplified geological classification developed for radon mapping in England and Wales (Miles and Appleton, 2005) was used in order to have an adequate number of soil samples on each PM class to facilitate mapping of spatial variation in element concentrations within the PM units.

Parent material polygons were subdivided into separate 200 m square polygons of the British National Grid using ESRI® ArcGIS geoprocessing tools to produce the shape files which store non-topological geometry and attribute information for the spatial features that form the basis for the production of the geochemical maps. Parent material codes and codes for the relevant 200 m grid square are attached to the locations of all soil samples. Parent material geochemical mapping was executed using an ArcGIS tool written in Vb.Net, as explained in Appleton and Adlam (2012). The optimum number of samples for calculating GM element concentrations for 200 m-PM polygons was not re-evaluated for this study, but hold out validation studies in England and Wales (Appleton et al., 2008) and Northern Ireland (unpublished data) indicated that the optimum number is between 4 and 7 for topsoils. For this study, GMs were calculated using data for the nearest 4 samples on the same PM apart from (a) three very minor PM units (calcrete, peat and the Upper Greensand) which comprise about 0.2% of the study area and which have no soil samples located on them and (b) polygons for which the average distance to the four samples required to calculate the GM was greater than 7000 m, which comprise about 4% of the study area. This largely affects sinuous polygon features, such as narrow alluvium areas with few 'local' soil samples along the upper reaches of tributaries to the river Thames. This approach was adopted to prevent excessive extrapolation of high element concentrations related to anthropogenic contamination, such as Pb. For both these sets of polygons, the GM was calculated from soil chemistry data for the nearest four samples, irrespective of PM.

3. Results and discussion

3.1 Summary statistics and ANOVA

Summary statistics (Table 3) were determined for the 8400 topsoils in the London region, subdivided into urban (n = 4710) and rural (n = 3690) domains. The urban domain was defined using the UK Ordnance Survey Strategi® 1:250,000 scale urban land use data (Figure 2), which includes areas mapped as 'Large urban areas' and 'Small urban areas'. In this vector dataset, an urban area is defined as containing a concentration of buildings and other structures. All areas out with those polygons were defined as 'rural'; it should be noted that this includes areas within central London which form large open-spaces, such as Richmond Park and Wimbledon Common in south-west London (Knights and Scheib, 2011) and differs somewhat from other definitions of 'rural'.

The proportion of the variance (R^2 (adjusted) in Table 4) explained by PM for the whole data set varies from 5% for Cd to 33% for Mn. When the data are subdivided into urban and rural domains, R^2 (adj) is moderate to high (24-37) and approximately the same in the two domains for those elements whose distribution is largely controlled by PM (for example, Al, Cs, Fe, Ga, K, Mg, Rb, Sc, Ti and V). This suggests that geochemical patterns related to PM which characterise rural domains are also detectable in the London urban domain, as has been previously observed in the Northampton area (Appleton and Adlam, 2012). For a number of elements whose distribution is influenced both by PM and anthropogenic factors, R^2 is substantially higher in the rural domain compared with the urban domain (e.g. As, Ca, Mo, Ni, Pb, Sb, Sr, Zn). ANOVA results for the whole of the dataset may be influenced by the unequal numbers of PMs in the urban and rural domains (15 and 19, respectively) so the ANOVA was repeated on a subset of the data (Figure 2 and Table 4) in which an equal number of PMs (13) are represented in both the rural and urban sample populations. The difference between R^2 for the urban and rural domains is generally slightly reduced but still

quite high (>10% difference) for As, Ca, Hf, I, Mn, Mo, Ni, Se, Sr, Zn and Zr which implies that the distribution of these elements has been impacted by anthropogenic factors in the urban domain. This is to be expected for As, Ca, Mo, Ni, Se, Sr and Zn which may be influenced by a range of anthropogenic factors including coal burning, spreading of ash from coal fires and from lime-based mortar and concrete dust but is rather surprising for elements such as Hf, I and Zr. In the topsoil subset representing equal number of PMs, R^2 is moderate to high (>20%) and similar in the urban and rural domains for Al, Ce, Cs, Fe, Ga, K, La, Mg, Nb, Nd, Rb, Sc, Th, Ti, and V (Table 4) implying that PM exerts a strong control on these elements in both domains. R^2 in both urban and rural domains is very low (5%) for Cd and only slightly higher in the rural domain compared with the urban domain for other elements strongly influenced by anthropogenic factors, such as Cu, Mo, Pb and Sn (Table 4).

3.2 Cluster analysis

Cluster analysis of variables in topsoil data for the current rural domain was used to identify groups of elements which share common characteristics and distributions (Table 5, Figure 3). Cluster 1 (Al, Cr, Cs, Fe, Ga, K, Mg, Nb, Rb, Sc, Th, Ti, V) elements all tend to be highest in clay-rich parent materials (Gault Formation, Thames Group (clay), Clay-with-flints, Glacial till, Head (clay-silt), alluvium and to a lesser extent Brickearth (e.g. Al in Figure 4) with the notable exception that K is not high in Clay-with-flints. This group of elements is correspondingly low in the other PMs. The strongest similarities in Cluster 2 (As, Br, Ge, Se and LoI) are (i) between Br, Se and LoI which reflects the well documented adsorption of Br and Se onto organic material and (ii) between As and Ge which may reflect both the geochemical association between these elements and their occurrence in coal ash from domestic fires that was historically disposed of in urban gardens and possibly also adsorbed onto organic material. As, Ge, and Se are slightly enriched in most of the clay PM's, although

As is not high in topsoils over the Gault Formation. Zr and Hf, probably in detrital zircon, are associated with Si (Cluster 3) in most of the arenaceous PMs including the Bagshot Formation, Bracklesham Group (sands, sand-silt), Thanet Sand Formation, Lower Greensand, and River Terrace Deposits (Supplementary Material Figure 1). Zr and Hf are both enriched also in Clay-with-flints. The inclusion of Ba in this cluster is more difficult to understand, especially as Ba is highest in clay-rich parent materials (Brickearth, Thames Group (clay), Head (clay-silt), Glacial till, Alluvium and Gault Formation).

Cluster 4 (Ca, Co, Mn, Ni, Sr and pH) reflects high concentrations of Ca and Sr associated with high pH in soils derived from the Grey Chalk and White Chalk and also higher than average Ca and Sr concentrations associated with clay PMs including the Gault, Thames Group (clay), Glacial Till, Alluvium and Head (clay-silt). Mn enrichment is common in residual soils developed over the Chalk; Co and Ni are associated with Mn due to adsorption effects.

The cluster 5 elements (Cd, Cu, Mo, P, Pb, Sb, Sn, Zn) do not demonstrate a consistent relationship with PM, having generally low to moderate R^2 values (Table 4) and their distribution will reflect a range of anthropogenic contamination factors including sewage disposal on agricultural land, fertilisers (Cd-P association), and historical use of lead in petrol. In the rural domain, mean Cu, Mo, P, Pb, and Zn are at a maximum in topsoils overlying alluvium. This is possibly caused by dispersion from present and historical industry which is frequently located on the Thames and Lee valley alluvium.

Cluster 6 (Ce, I, La, Nd, Sm, U, Y, Yb) reflects high concentrations of this group of elements in soils over chalk, glacial till (which is predominantly chalky) and Clay-with-flints, which overlie the chalk and to a lesser extent with clay PMs including the Gault, Thames Group (clay), head (clay-silt), glacial till, Brickearth and alluvium in the case of Ce, La (Figure 5), and Y. There is a strong relationship between the individual rare earth elements (especially

Ce, La, Nd, and Y) due to their similar chemical properties. Higher concentrations of I are particularly associated with the Chalk, Clay-with-flints developed over chalk and the predominantly chalky Glacial Till due to the strong chemical affinity of I with soils developed over carbonate rocks. Uranium is highest in Clay-with-flints, probably associated with detrital minerals such as zircon.

3.3 PM soil chemistry mapping

Interpolation of the topsoil data to 200 m-PM polygons serves to constrain the 'leakage' of high element concentrations, such as Ca in the Chalk into areas underlain by adjacent PMs with lower Ca concentrations, such as the Clay-with-flints. In the south east sector of the London region, very low Ca also characterises the sands of the Thanet Sand Formation and Lower Greensand Group. The PM mapping method is particularly effective where the distributions of elevated concentrations are spatially complex, since they relate to the convoluted outcrop pattern (Figure 6 and Supplementary Material Figure 2). As a consequence, errors at PM boundaries are likely to be reduced when estimates of element concentrations are derived within delineations of the PM mapping units. The proportion of the variance explained by PM for Al, Ca and Si is 33, 43 and 51%, respectively in the south-east sector.

The description of the main characteristics of the PM geochemical mapping that follows is based on selecting the dominant one or more elements from each of the main clusters described above.

There is a close association of higher Al (Figure 7) and Th (Supplementary Material Figure 3) concentrations with the London Clay (Thames Group clays), Gault Formation (clay), clay-

with-flints, alluvium (clay, silt and subsidiary sand and gravel), and glacial till. As would be expected, lower concentrations are associated with the predominantly arenaceous River Terrace deposits, Plateau Gravels, Bagshot Formation (sands), Thanet Sand Formation and Lower Greensand Group (sands) as well as the Chalk.

Calcium has a much higher proportion of its variance controlled by PM in the rural domain (Table 4 and Supplementary Material Figure 4) where very high concentrations of Ca are associated with the chalk and higher than average Ca associated with clay PMs including the Gault, Thames Group (clay), Glacial Till and alluvium and Head (clay-silt). Enrichment of Ca in topsoils of the urban domain (median 10373 mg kg^{-1}) compared with the rural domain (median 6255 mg kg^{-1}) may be caused by cement and lime-based mortar dust being incorporated into topsoils as a result of the widespread destruction of buildings across large sectors of the London urban domain especially during the period 1940-41.

The PM map for Si (Figure 8) illustrates the very low concentrations in topsoils over the Grey and White Chalk and particularly high concentrations in most of the sand PMs including the Lower Greensand, Thanet Sand Formation, Bagshot Formation, Bracklesham Group and River Terrace Deposits. Relatively high Si in soil over some areas mapped as Thames Group (clay) may reflect the occurrence of a thin layer of River Terrace Deposits at the surface that was considered to be too thin (<1 m) to be mapped as a superficial deposit.

The distribution of Mn (Supplementary Material Figure 5) and La (Figure 9) illustrates high concentrations in topsoils over chalk, glacial till (which is predominantly chalky) and the Clay-with-flints, which overlie the chalk and also, but to a lesser extent with clay PMs including the Gault, Thames Group (clay), Head (clay-silt), Glacial Till, Brickearth and Alluvium in the case of La (Figure 9).

Although As (Supplementary Material Figure 6) is slightly enriched in most of the clay PM's Clay-with-flints, Thames Group (clay), and alluvium) and also with the Lower Greensand in

the rural areas, the dominant association is with anthropogenic contamination in the urban domain. This is most pronounced with alluvium in the Lee and Thames valleys, partly due to co-location and concentration of industry on this PM and partly due to dispersion and concentration in sediments in the lower Thames floodplain, which, like many estuarine soils (Fletcher et al., 1994a, b) are characterised by relatively high concentrations of a range of industrially derived elements including As, Cd, Pb, Hg and Se.

Iodine and Br have a similar distributions being particularly associated with the Chalk, Clay-with-flints developed over the chalk and the predominantly chalky Glacial Till (Figure 10) due to the strong chemical affinity of I in particular, with soils developed over carbonate rocks (Rawlins et al., 2012). Relatively high concentrations of both I and Br in the topsoils developed over alluvium in the lower section of the Thames valley probably reflect the impact of salt water, whilst higher I and Br associated both with alluvium and plateau gravels in the west of the London region suggest redistribution of carbonate from the Chalk outcrops, which is confirmed by the Ca map (Supplementary Material Figure 4). The spatial distributions of Br and LoI are more complex, with higher concentrations of both being associated with the Thames Group clays and with alluvium in the lower Thames valley, sectors of the Lee valley as well as with alluvium in some of the smaller valleys. The strong correlation between LoI and Br and the lack of correlation between LoI with pH or Ca suggests that most of the LoI is related to organic material and very little to soil carbonate.

Topsoil Pb (Supplementary Material Figure 7) is dominantly controlled by anthropogenic factors, being high throughout the urban domain. High Pb is also associated with 'rural' topsoils overlying alluvium due to the impact of industrial development in the Thames and Lee valleys but relatively low over some of the major parks in south west London, which have not been impacted by significant urban development throughout the 200-300 year history of London (Knights and Scheib, 2011).

The Zr map (Supplementary Material Figure 8) illustrates that the highest concentrations occur over Clay-with-flints (median 439 mg kg^{-1}), and the Bracklesham Group (sand; median 368 mg kg^{-1}) and also the Bagshot Formation (sand; median 386 mg kg^{-1}). This supports the genesis of the Clay-with-flints by weathering and solifluction of original Palaeogene deposits. Part of the Zr in the Clay-with-flints may have been derived by weathering of the underlying White Chalk (median 289 mg kg^{-1}). An isolated patch of high Zr and Hf is associated with the outlier of Bagshot Formation in the Hampstead Heath area (Supplementary Material Figure 8), although the anomaly extends more than a kilometre laterally over the surrounding Thames Group (clay) suggesting either the outcrop of the Bagshot Formation is larger than mapped or that a thin cover of residual zircon enriched sandy soils derived from the Bagshot Formation occur over the Thames Group (clay). The small proportion of the variance of Hf and Zr controlled by PM in the urban domain, which lies mainly within the Greater London Authority (GLA) boundary (Supplementary Material Figure 8), may reflect the broadly similar concentrations of these elements in the main geological units found within the GLA (Figure 1).

4. Conclusions

This study has established that geogenic signatures which dominate soil chemistry in rural domains can also be readily detected in the London urban region. ANOVA of the soil data for the London region indicates that 25-33% of the variance of Al, Ce, Cs, Ga, K, La, Mg, Mn, Nb, Nd, Rb, Ti, V and Y is explained by soil PM (surface geology), and a slightly lesser proportion (19-25%) of the variance for Ca, Co, Fe, I, Ni, Sc, Sr and Th. The variance of some other elements is influenced by a mixture of geogenic and anthropogenic controls, including Ba, Pb, Sb, Sn and Zn for which PM controls 12-16% of the variance. Geogenic soil chemistry patterns observed in the rural areas surrounding London can be quite clearly followed into and through the London urban area, confirming that the spatial patterns of a

range of elements primarily controlled by PM has not been destroyed even in a major urban centre that has been subjected to intensive urban development, destruction and redevelopment during the last 200-300 years.

Acknowledgements

This paper is published with the permission of the Director of the British Geological Survey (Natural Environment Research Council). G-BASE is a BGS national capability project to map the surface environment of Great Britain and Northern Ireland. The London region soils were collected as part of the London Earth Project and we acknowledge that the authors are reporting work based on a sampling project led by Cathy Scheib, Kate Knights, Andreas Scheib, Bob Lister, Mick Strutt, Jenny Bearcock and Paul Everett. Sample preparation and analysis was done in the BGS laboratories (Keyworth) and the efforts of Mark Allen, Kevin Barker, Charles Brettle, John Wheeler, Les Merry, Mark Ingham, Charles Gowing, Heather Harrison, Simon Carter, Neil Eatherington and Leian Grimsley are acknowledged. Bob Lister and Kay Green were responsible for the data quality control and levelling. The work would not have been possible without the efforts of the student sampling teams who collected the thousands of soils reported here – their contribution is gratefully acknowledged. Finally we thank Kate Royse and Neil Breward (BGS), C. Reimann, M. Birke and one anonymous reviewer for suggesting useful improvements to earlier versions of this paper.

References

Aldiss, D.T., 2012. The stratigraphical framework for the Palaeogene successions of the London Basin, UK. British Geological Survey Open Report (OR/12/004) http://nora.nerc.ac.uk/17121/ Last accessed 11.02.2013

Allen, M.A., Cave, M.R., Chenery, S.R.N., Gowing, C.J.B., Reeder, S., 2011. Sample Preparation and Inorganic Analysis for Urban Geochemical Survey Soil and Sediment Samples, in: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 28-46.

Appleton, J.D., Adlam, K.A.M., 2012. Geogenic control on soil chemistry in urban areas: A novel method for urban geochemical mapping using parent material classified data. Appl. Geochem. 27, 161-170.

Appleton, J.D., Rawlins, B.G., Thornton, I., 2008. National-scale estimation of potentially harmful element ambient background concentrations in topsoil using parent material classified soil: stream-sediment relationships. Appl. Geochem. 23, 2596-2611.

BGS 2011. London Earth. http://www.bgs.ac.uk/gbase/londonearth.html Last accessed 11.02.2013

De Vos, W., Tarvainen, T. (Editors), 2006. Geochemical Atlas of Europe. Part 2. Interpretation of Geochemical Maps, Additional Tables, Figures, Maps, and Related Publications. Geological Survey of Finland, Espoo, Finland.

Ellison, R.A., Woods, M.A., Allen, D.J., Forster, A., Pharaoh, T.C., King, C., 2004. Geology of London. Memoir of the British Geological Survey, Sheets 256 (North London), 257 (Romford), 270 (South London) and 271 (Dartford) (England and Wales). British Geological Survey, Keyworth, Nottingham.

Fletcher, C.A., Bubb, J.M., Lester, J.N., 1994a. Magnitude and Distribution of Contaminants in Salt-Marsh Sediments of the Essex Coast, Uk .1. Topographical, Physical and Chemical Characteristics. Sci. Total Environ. 155, 31-45.

Fletcher, C.A., Bubb, J.M., Lester, J.N., 1994b. Magnitude and Distribution of Anthropogenic Contaminants in Salt-Marsh Sediments of the Essex Coast, UK .2. Selected Metals and Metalloids. Sci. Total Environ. 155, 47-59.

Flight, D.M.A., Scheib, A.J., 2011. Soil Geochemical Baselines in UK Urban centres: The G-BASE Project, in: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 186-206.

Hopson, P.M., Wilkinson, I.P., Woods, M.A., 2008. A stratigraphical framework for the Lower Cretaceous of England. Keyworth, Nottingham, British Geological Survey Research Report (RR/08/003), 77pp. http://nora.nerc.ac.uk/3236/ Last accessed 11.02.2013

Johnson, C.C., Breward, N., Ander, E.L., Ault, L., 2005. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. Geochemistry: Exploration-Environment-Analysis, 5, 1-13.

Johnson, C.C., 2011. Understanding the Quality of Chemical Data from the Urban Environment – Part 1: Quality Control Procedures, in: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 61-76.

Johnson C.C., Demetriades A., Locutura J., Ottesen R.T. (Editors), 2011. Mapping the Chemical Environment of Urban Areas. John Wiley & Sons Ltd., Chichester, UK.

Johnson C.C., Ander E.L., Cave, M.R., Palumbo-Roe, B. 2012. Normal background concentrations (NBCs) of contaminants in English soils: Final project report. BGS Commissioned Report CR/12/035.British Geological Survey, Keyworth. http://nora.nerc.ac.uk/19946 Last accessed 11.02.2013

Knights, K., Scheib, C., 2011. Examining the soil chemistry of London's parklands. [Poster] In: Cities, catchments and coasts: applied geoscience for decision-making in London and the Thames Basin, London, UK, 13 May 2011. http://nora.nerc.ac.uk/14272/ Last accessed 11.02.2013

Miles, J.C.H., Appleton, J.D., 2005. Mapping variation in radon potential both between and within geological units. J Radiol Prot 25, 257-276.

Rawlins, B.G., Webster, R., Lister, T.R., 2003. The influence of parent material on topsoil geochemistry in eastern England. Earth Surf. Process. Landf. 28, 1839-1409.

Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C., Carter, S. 2012. The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth. http://nora.nerc.ac.uk/18016/ Last accessed 11.02.2013

Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical Data Analysis Explained. Applied Environmental Statistics with R. Wiley, Chichester.

Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P. (Editors), 2013. Chemistry of Europe's agricultural soils. Geologisches Jahrbuch (Reihe B) (available: 5 December 2013).

Royse, K.R., de Freitas, M., Burgess, W.G., Cosgrove, J., Ghail, R.C., Gibbard, P., King, C., Lawrence, U., Mortimore, R.N., Owen, H., Skipper, J., 2012. Geology of London, UK. P Geologist Assoc 123, 22-45.

Salminen, R., Batista, M.J., Bidovec, M., Demetriades, A., De Vivo, B., De Vos, W., Duris, M., Gilucis, A., Gregorauskiene, V., Halamic, J., Heitzman, P., Lima, A., Jordan, G., Klaver, G., Klein, P., Lis, J., Locotura, J., Marsina, K., Mazreku, A., O'Connor, P.J., Olsson, S.A., Ottesen, T., Petersell, V., Plant, J.A., Reeder, S., Salpeteur, I., Sandstrom, H., Siewers, U., Steenfelt, A. and Tarvainen, T., 2005. Geochemical atlas of Europe. Part 1, Background information, methodology and maps. Geological Survey of Finland, Espoo, Finland.

Scheib, A., Flight, D., Lister, B., Scheib, C., 2011 London Earth: anthropogenic and geological controls on the soil chemistry of the UK's largest city. [Unpublished poster presentation] In: 25th International Applied Geochemistry Symposium, Rovaniemi, Finland, 22-26 Aug 2011. Available at http://nora.nerc.ac.uk/15033/ Last accessed 11.02.2013

Table 1 Summary of bedrock geology of the London region (after Aldiss, 2012; Ellison et al., 2004; Hopson et al., 2008; Royse et al., 2012; units highlighted in **bold** text used in Figures 1, 3-5)

PERIOD	GROUP	FORMATION	THICKNESS (M)
PALEOGENE	Bracklesham	**Bracklesham Group(sand):** Camberley Sand Formation (sands) and Windlesham Formation (sands) **Bracklesham Group(sand-silt):** Windlesham Formation: sand, silt and minor clay	na
		Bagshot Formation: sand, fine-grained with thin clay beds	10-25
	Thames	**Thames Group (clay)**: London Clay Formation: clay, silty; fine sand clay at base. Claygate member: interbedded sand and clay at top	90-130
		Thames Group (sand): Harwich Formation: sand, clayey fine grained sand and pebble beds	0-10
	Lambeth	Reading & Woolwich Formations: clay mottled with fine-grained sand, laminated clay, and shelly clay Upnor Formation: sandstone with flint pebble beds	10-20
	Montrose	**Thanet Sand Formation**: sand, fine-grained	0-30
CRETACEOUS	**White Chalk**	Seaford, Newhaven, Lewes Nodular, New Pit, and Holywell Nodular Chalk Formations: white chalk, soft to hard and nodular, sometimes with flint courses	Up to 180
	Grey Chalk	Zig Zag and West Melbury Chalk Formations: chalk, pale grey with thin marls; glauconitic at the base	40-80
	Selborne	**Upper Greensand** Formation: sand fine-grained, glauconitic	Up to 17
		Gault Formation: clay, silty	50-70
	Lower Greensand	Folkestone Formation : Sandstone, fine to medium-grained	60
		Sandgate Formation: sandstone Hythe Formation: mudstone	34
	Wealden	Weald Clay Formation: mudstone Hastings Beds: sandstone and mudstone	Up to 150

Table 2 Quaternary superficial deposits in the London region (after Ellison et al., 2004 and Royse et al., 2012)

Quaternary Deposit	Lithology
Head (silt-clay)	Silt-clay solifluction deposits largely derived from London Clay Formation
Head (gravel-sand)	Gravel-sand solifluction deposits largely derived from Plateau Gravels and River Terrace Deposits
Alluvium	Silty-clay and clayey-silt with subsidiary sands
Brickearth	Fine-grained sand, silt and clayey silt (loessic deposits) with basal gravels (solifluction and fluvial)
River Terrace Deposits	Gravel, gravelly-sand and sand with flint, vein quartz and quartzite clasts
Glacial till	Formerly known as "Chalky Boulder Clay": pebbly to boulder-rich clay with chalk and flint clasts (subsidiary Triassic sandstone, vein quartz and quartzite clasts)
Plateau gravels	Quart, quartzite and flint pebbles; mainly form hill-top caps
Clay-with-flints	Clay, silty-clay and sandy-clay with flint clasts stained with manganese formed by weathering and solifluction of Palaeogene cover and dissolution of Chalk

Table 3 Summary statistics for topsoil samples from the London region rural and urban domains

	Urban Domain (n = 4710; n LOI and pH = 4684)					Rural Domain (n = 3690;n LOI and pH =3245)				
Variable	Mean	Min.	Med.	Max.	Skew.	Mean	Min.	Med.	Max.	Skew.
Al	40878	6877	38617	109503	0.7	44518	4232	43378	134895	0.2
As	17.6	1.2	15.9	161	5.5	14.6	1.2	13.6	126	5.9
Ba	418	159	389	3475	6.5	354	139	349	2333	6.6
Br	13.0	2.2	12.1	93	2.7	12.2	1.1	10.6	241	9.3
Ca	15562	786	10373	303319	5.5	17547	143	6255	350060	4.7
Cd	1.0	0.1	0.6	110	25.5	0.9	0.05	0.5	165	27.5
Ce	50	20.0	48.7	238	2.1	54	16	53	170	1.5
Co	12.1	1.0	11.5	52	1.7	12.2	0.05	11.3	85	2.7
Cr	75	21	71	532	4.3	81	9	76	2094	17.9
Cs	3	1	3	11	1.1	3.4	1	3	11	0.9
Cu	78	5	52	4577	15.6	46	3	25	5326	25.0
Fe	27976	3812	26793	107507	1.1	26855	909	26229	109036	0.7
Ga	10.9	1	10.4	24	0.9	11.4	3.1	11.1	44	0.8
Ge	2.1	0.1	1.7	29	5.1	1.3	0.05	1.1	39	12.5
Hf	7.3	2.2	7	37	2.4	8.7	1.1	8.3	41	1.6
I	3.4	0.2	2.9	48	6.4	4.6	0.1	3.5	80	6.4
K	11370	1577	10791	27642	0.6	12108	996	11372	28804	0.3
La	24	5	23	130	2.2	28	3	27	134	2.1
LOI	7.7	0.7	7.3	32	1.6	7.6	1.2	6.9	72	4.7
Mg	5097	603	4825	24124	1.7	5304	603	4825	27743	1.6
Mn	462	31	434	5398	5.9	582	15	457	4732	2.9
Mo	1.9	0.1	1.6	41	9.1	1.8	0.05	1.1	561	36.1
Nb	12.5	5.7	12.2	22	0.5	13.7	4.7	13.6	147	11.7
Nd	22	2	21.1	115	2.0	24	0.1	22.8	173	2.9
Ni	28	3	26.1	358	5.8	26	1	22.5	506	9.9
P	1685	262	1484	14270	4.9	1246	44	1091	19594	8.6
Pb	344	17	220	25206	22.0	132	10	73	13557	25.5
pH	6.4	3.1	6.6	8	-1.1	5.9	2.8	6.21	8	-0.6
Rb	58	12.7	54.3	149	0.8	64	6.5	61.4	157	0.4
Sb	5.7	0.3	3.5	435	20.1	2.9	0.1	1.4	612	28.7
Sc	8.1	0.05	7.8	23	0.6	8.0	0.05	7.8	36	0.5
Se	0.7	0.05	0.6	16	12.0	0.6	0.05	0.5	20	14.2
Si	309230	35492	310555	467000	-0.7	308101	21482	312890	467000	-1.1
Sm	3.5	0.05	3.4	21	1.0	3.9	0.05	3.6	39	3.0
Sn	28	1.5	16.5	1042	8.6	12.8	0.9	6.1	630	10.7
Sr	85	19	78	601	2.8	79	11	68	576	3.4
Th	6.8	1.1	6.5	457	62.1	7.2	0.05	7.4	23	-0.2
Ti	3406	1249	3265	6847	0.6	3745	1048	3708	7069	0.1
U	1.6	0.05	1.6	6	-0.1	2.0	0.05	2.1	11	0.6
V	82	17	75	302	1.1	81	9	75	531	1.6
Y	21	4.9	20.2	112	3.5	23	3.4	21.5	158	3.7
Yb	1.9	0.05	1.8	9	0.5	2.1	0.05	2	13	1.3
Zn	247	11	181	8801	10.1	129	1	86	10095	22.8
Zr	284	56	272	1446	2.6	329	35	317	1488	1.5

All concentrations in mg kg^{-1} except LOI (% loss on ignition) and pH

Table 4 Analysis of variance: topsoil chemistry for London region rural and urban domains

	All of region				Subset (balanced geology)		
Domain	All samples	Urban	Rural		Urban	Rural	
N	8400	4710	3690		4679	3061	
DF	18	14	18		12	12	
	R^2(adj)	R^2(adj)	R^2(adj)	Rural - Urban*	R^2(adj)	R^2(adj)	Rural – Urban*
Al	30	30	33	3	30	30	0
As	16	9	22	13	9	19	10
Ba	15	8	19	12	7	15	8
Br	6	3	12	10	2	10	8
Ca	24	17	34	17	17	32	14
Cd	5	2	9	7	1	7	6
Ce	26	21	31	10	20	24	4
Co	21	14	28	15	14	22	8
Cr	16	17	17	0	17	15	-2
Cs	27	24	31	7	24	27	3
Cu	16	8	19	11	8	16	8
Fe	22	24	25	1	21	25	4
Ga	27	26	30	4	26	28	2
Ge	9	5	9	4	5	8	4
Hf	13	6	19	12	6	18	12
I	25	14	30	16	14	26	12
K	31	29	37	8	29	36	7
La	30	23	36	13	22	29	7
LOI	12	5	13	8	5	11	6
Mg	28	30	27	-3	31	28	-3
Mn	33	18	42	24	17	33	16
Mo	15	6	20	15	5	16	11
Nb	28	25	33	9	24	30	6
Nd	26	20	32	12	20	26	6
Ni	20	12	29	17	12	23	11
P	13	7	17	9	7	14	7
Pb	16	5	18	13	5	11	7
pH	13	7	22	15	7	18	11
Rb	29	27	33	5	27	29	2
Sb	16	5	19	14	5	13	8
Sc	23	24	25	1	24	24	1
Se	12	6	20	14	6	17	12
Si	19	15	23	8	15	21	7
Sm	7	5	10	5	5	8	4
Sn	16	8	19	11	7	13	6
Sr	20	13	28	15	12	23	11
Th	25	21	31	10	21	23	2
Ti	29	28	36	8	27	32	5
U	9	5	13	7	11	11	0
V	30	27	34	6	27	31	3
Y	26	18	33	16	17	27	10
Yb	12	7	13	6	7	11	5
Zn	14	4	21	17	4	17	13
Zr	13	7	18	12	6	17	11
n (pH LOI)	7928	4683	3245		4608	2839	

* R^2 (Rural) – R^2 (Urban)

Table 5 Results of cluster analysis of rural domain centred logratio transformed topsoil data (n 3245; Ward linkage, Correlation Coefficient Distance; final partition similarity level = 20)

Cluster No.	Variables
1	Al, Cr, Cs, Fe, Ga, K, Mg, Nb, Rb, Sc, Th, Ti, V
2	As Br Ge Se LoI
3	Ba Hf Si Zr
4	Ca Co Mn Ni Sr pH
5	Cd Cu Mo P Pb Sb Sn Zn
6	Ce I La Nd Sm U Y Yb

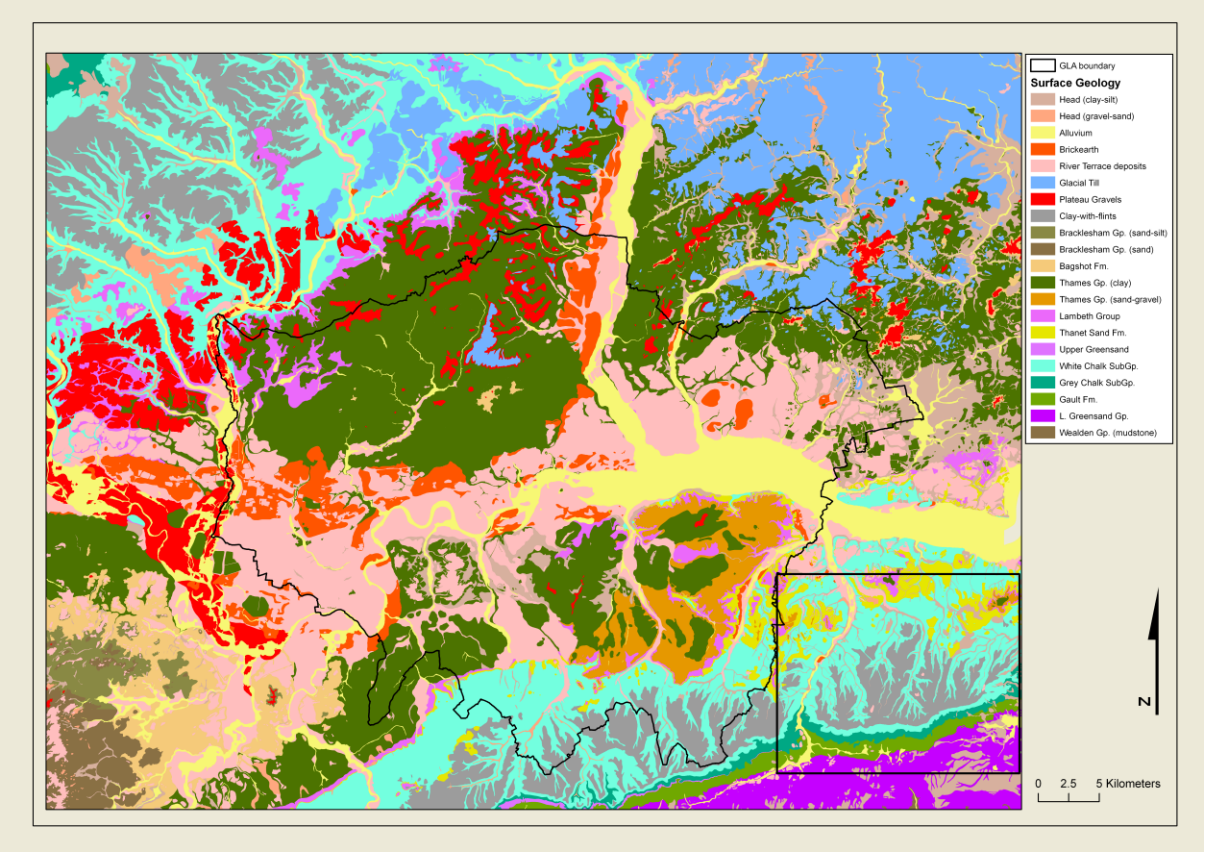

Figure 1 Simplified Soil Parent Material map of the London region showing the boundary of the urban area (GLA = Greater London Authority; box in SE corner indicates extent of Figure 6)

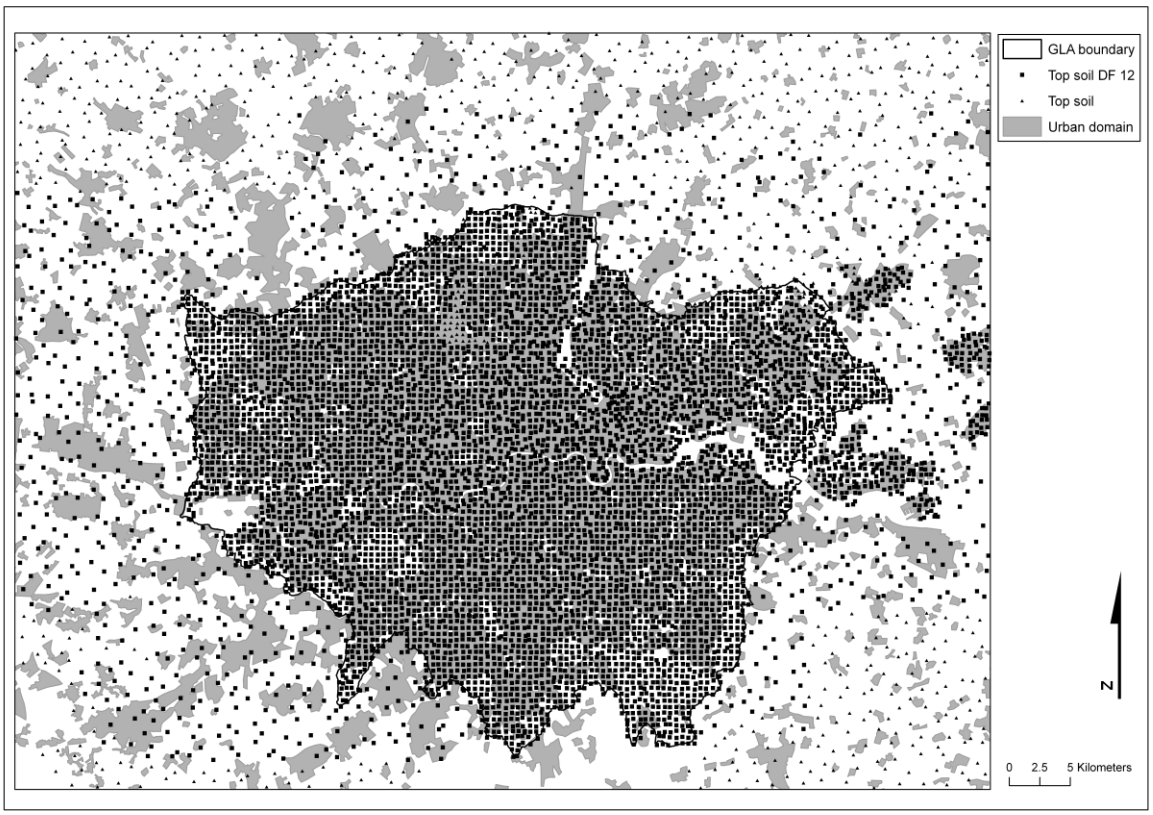

Figure 2 Location of top soil samples and extent of modern (2011) urban domain in the London Region (urban domain derived from OS Strategi® data)

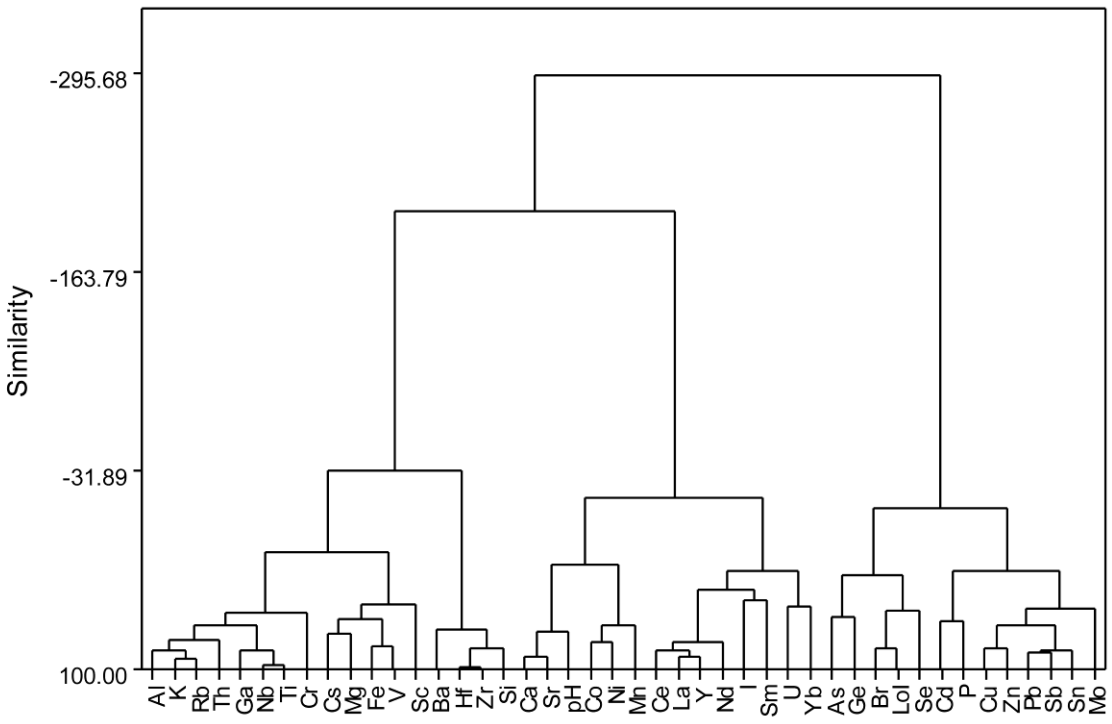

Figure 3 Cluster analysis dendrogram for rural domain soil chemistry (n = 3245)

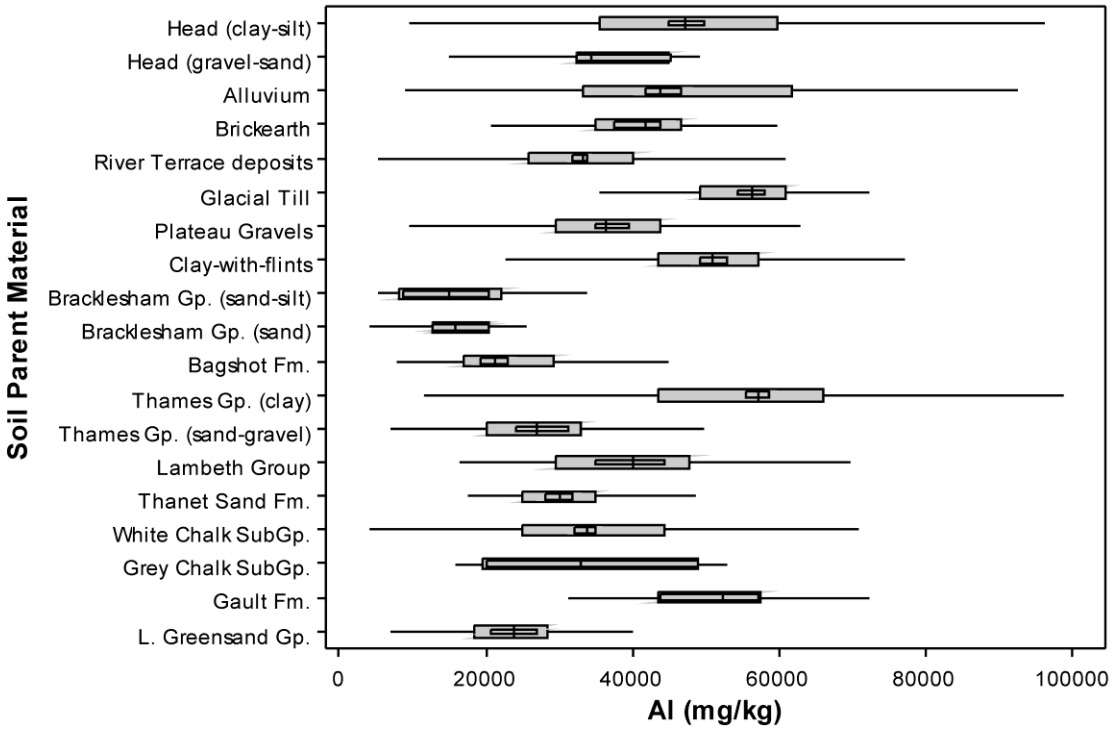

Figure 4 Boxplot of Al in topsoils from rural domain (n = 3245; Inner box = median confidence interval; outer box = interquartile range)

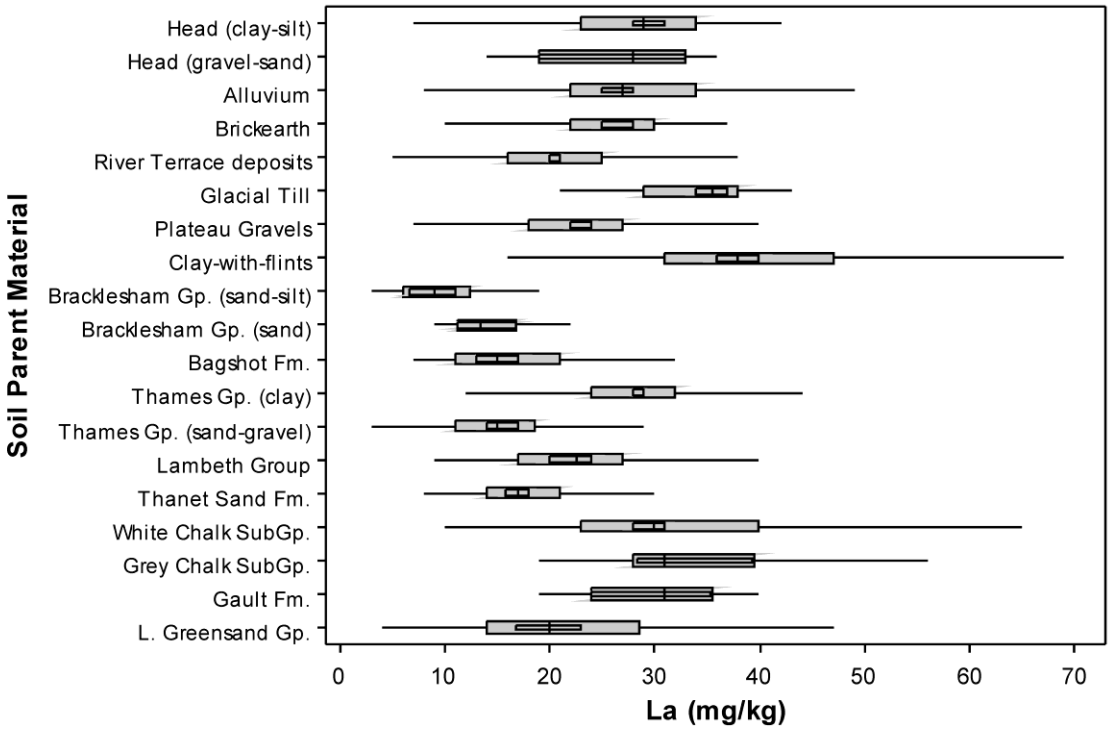

Figure 5 Boxplot of La in topsoils from rural domain (n = 3245; Inner box = median confidence interval; outer box = interquartile range)

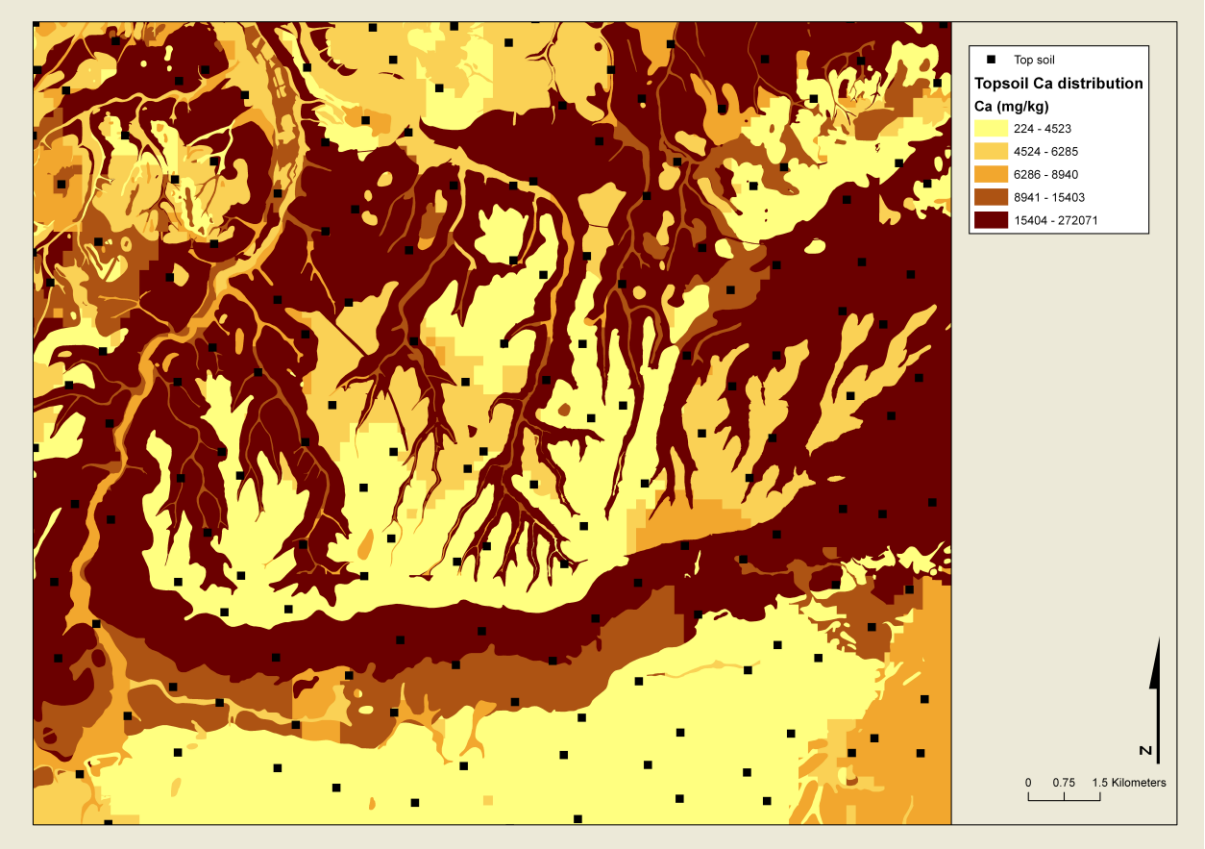

Figure 6 Topsoil Ca (■ = topsoil sample site) interpolated to 200 m-PM grid for the south east sector of the London region (see Figure 1 and Supplementary Material Figure 2 for soil PM map based on geology

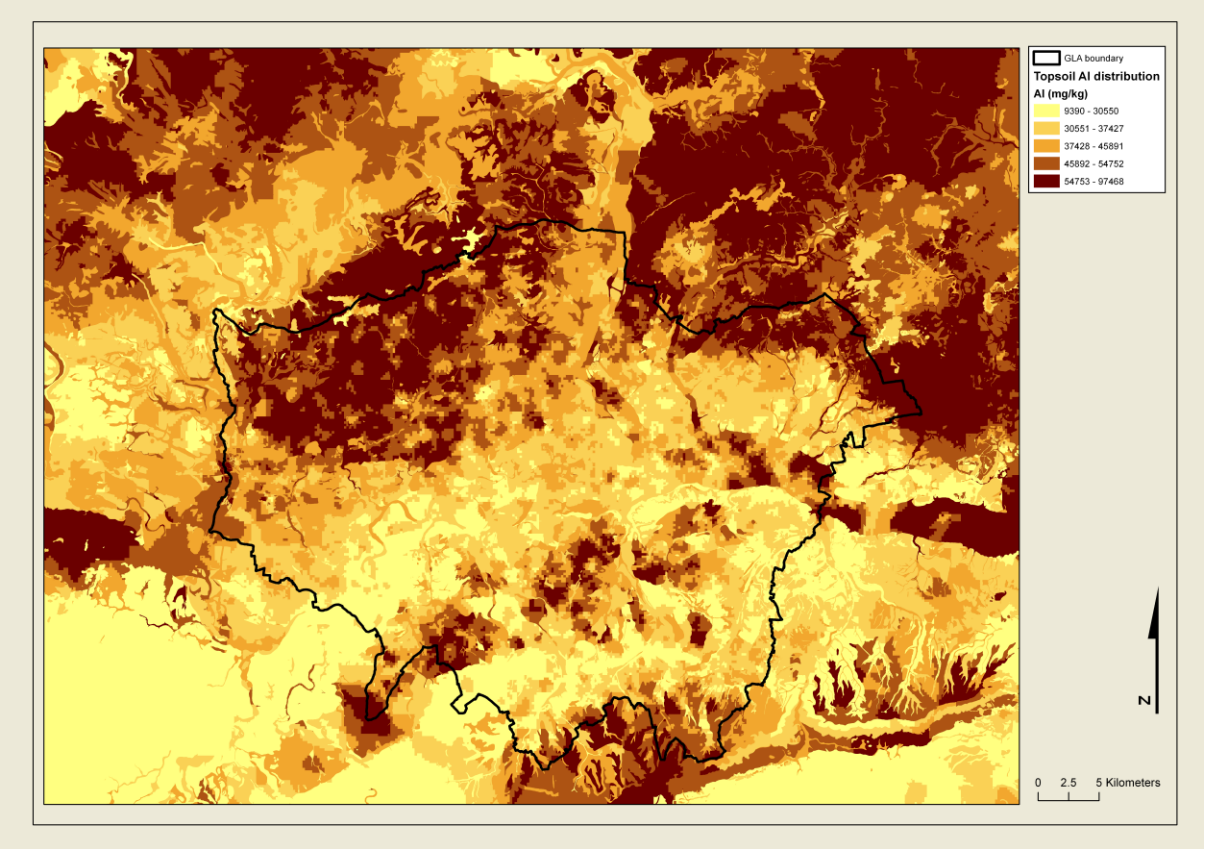

Figure 7 Topsoil Al interpolated to 200 m-PM grid

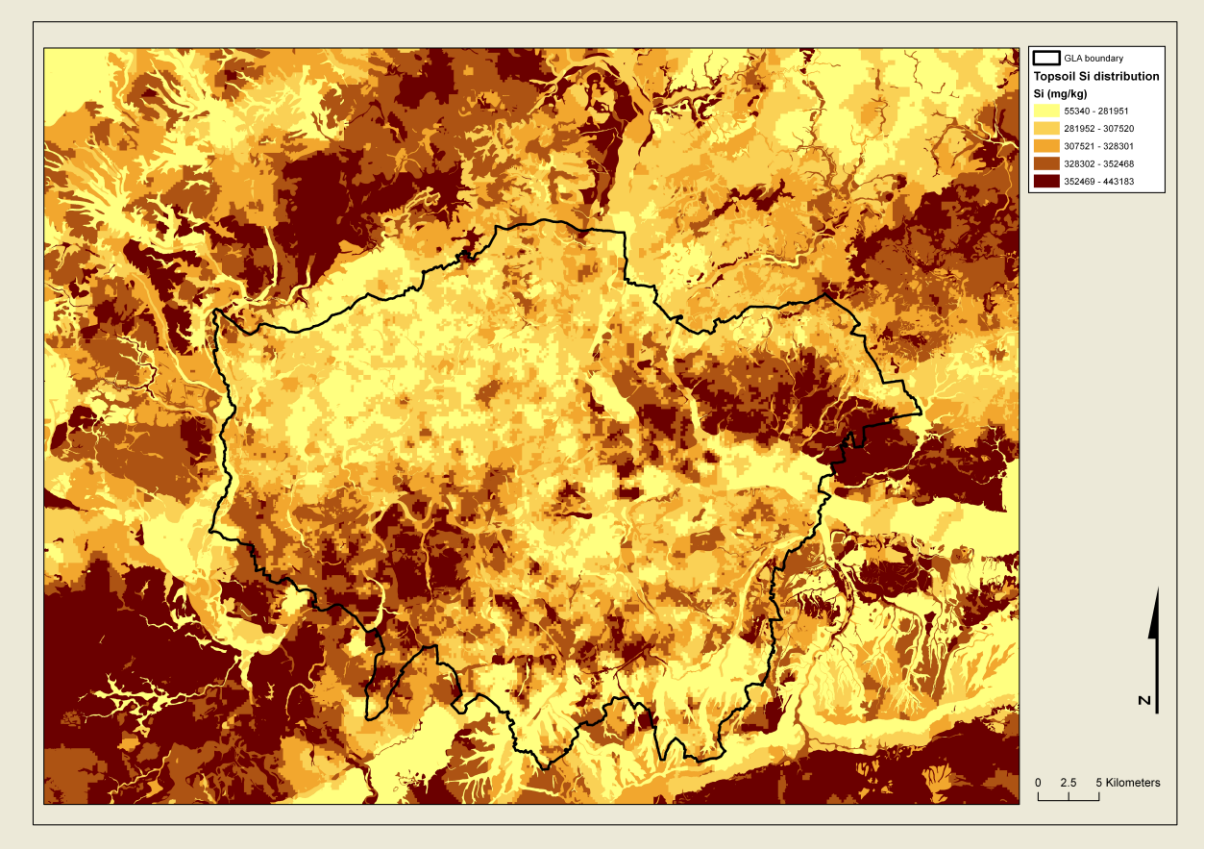

Figure 8 Topsoil Si interpolated to 200 m-PM grid

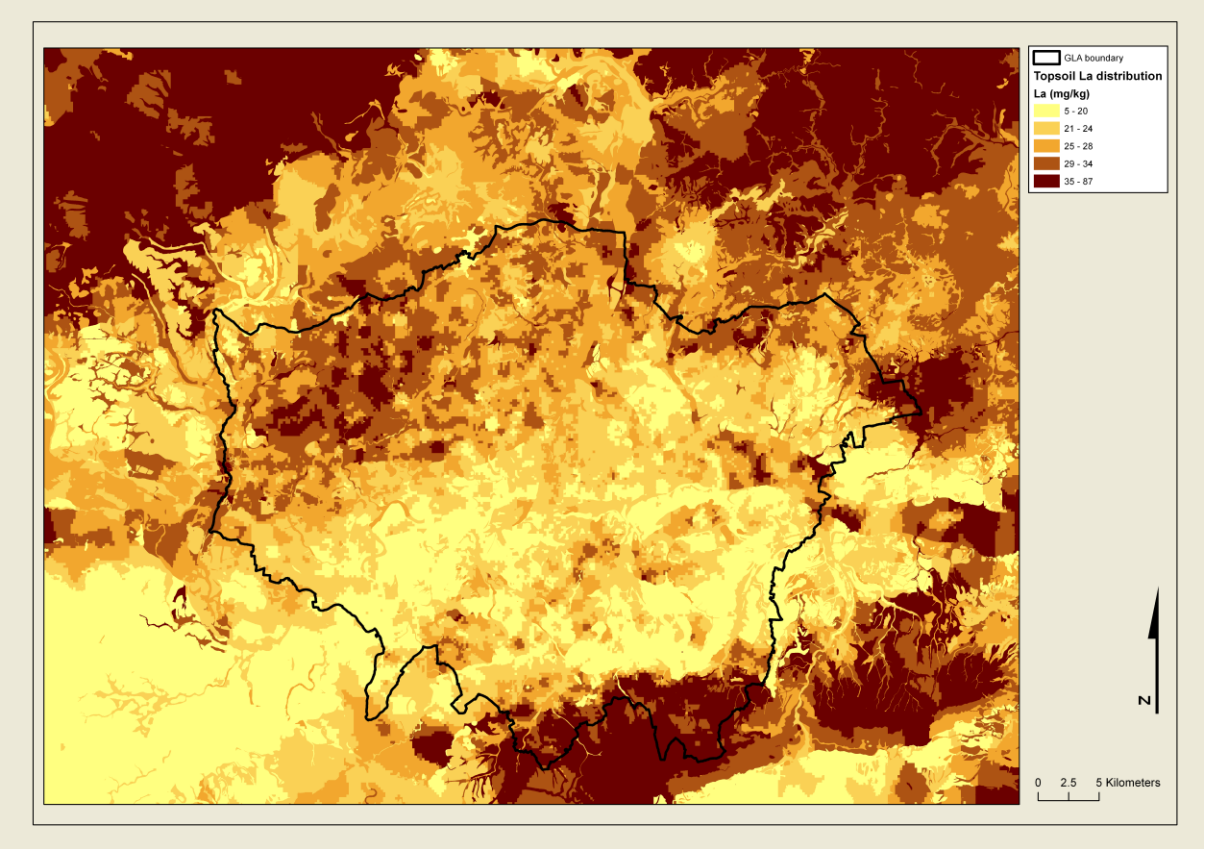

Figure 9 Topsoil La interpolated to 200 m-PM grid

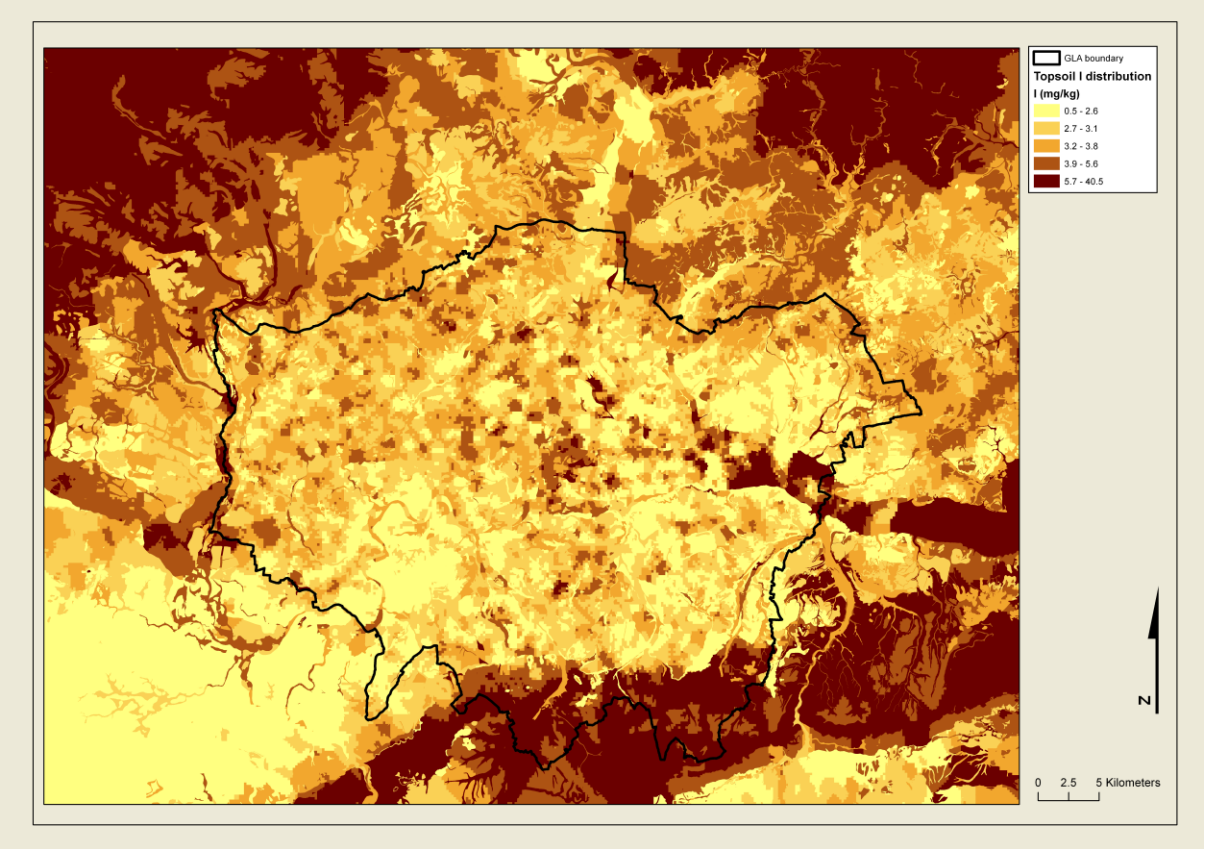

Figure 10 Topsoil I interpolated to 200 m-PM grid